

台北市各年收入族群

50萬以下 / 50-75萬 / 75-100萬 / 100-125萬 / 125-150萬 /

150-175萬 / 175-200萬 / 200萬以上

報告者:





# **目**錄CONTENT

Ⅲ 01 選取資料動機

Data selection and motivation

📶 02 圖表說明

**Graphics** presenting

🥃 03 問題與解決

Problems and solutions

💥 04 總結與感想

Conclusions



Data selection and motivation



# Ⅲ 01 選取資料動機



#### 為何選消費主題?

曾從事媒體教育產業行銷企劃人員,本身也很喜歡購物消費,想了解台北市各收入族群怎麼花錢及疫情後的消費狀況。



#### 資料來源

• 政府資料開放平臺 https://data.gov.tw/



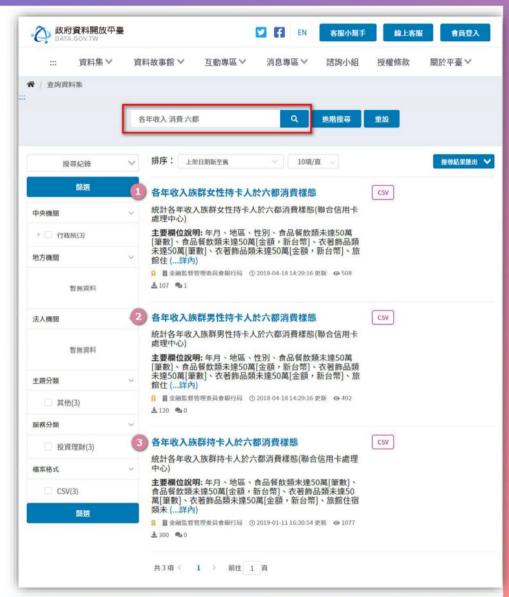
#### 檔案内容

- · CSV【各年收入族群持卡人於六都消費樣態】
- · JSON【各年收入族群持卡人於六都消費樣態】

# Ⅲ 01 選取資料動機









# 01 選取資料動機 - 觀察資料 CSV檔

01食品餐飲類

02衣著飾品類

03旅館住宿類

04交通類

05文教康樂

06百貨類



年收入 50萬以下

年收入 50-75萬

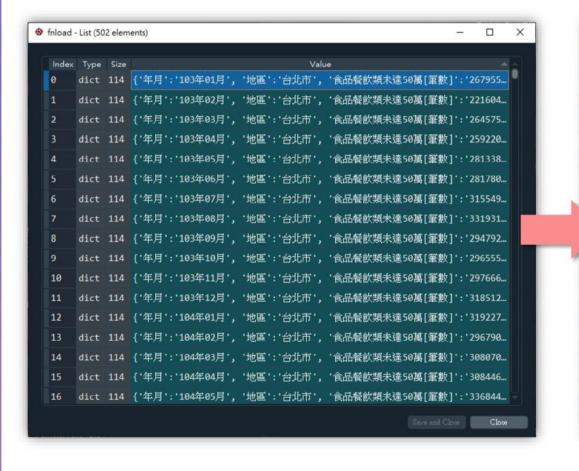
103年01月至109年11月

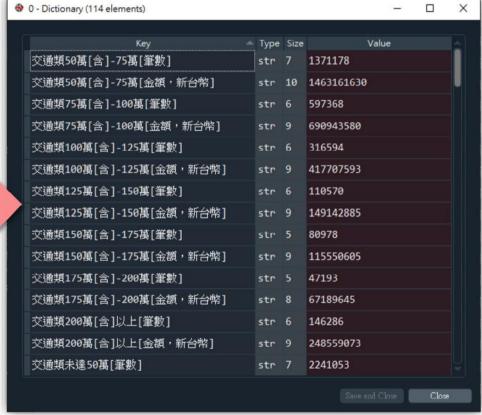
台北市等六都





## 01 選取資料動機 - 觀察資料 JSON檔



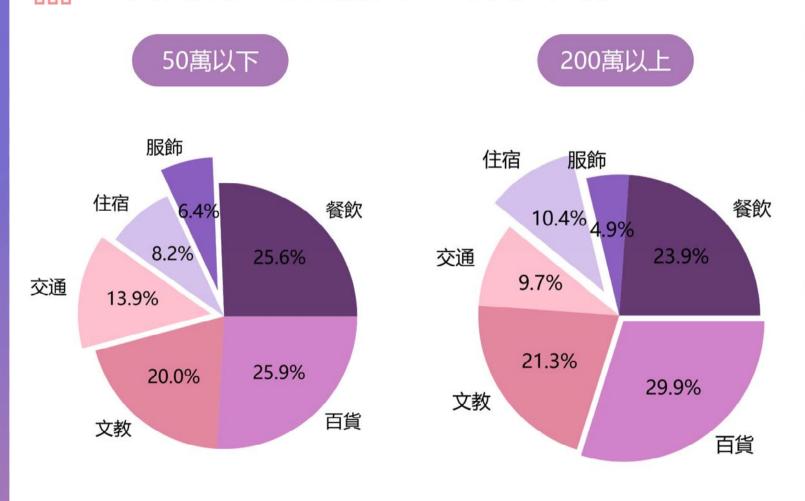


# **』** 02 圖表說明

**Graphics presenting** 



# | 02 圖表說明 – 消費輪廓 (109年度台北市)



#### 製圖說明

使用圓餅圖,並找適合的顏色之16進位法來上色。

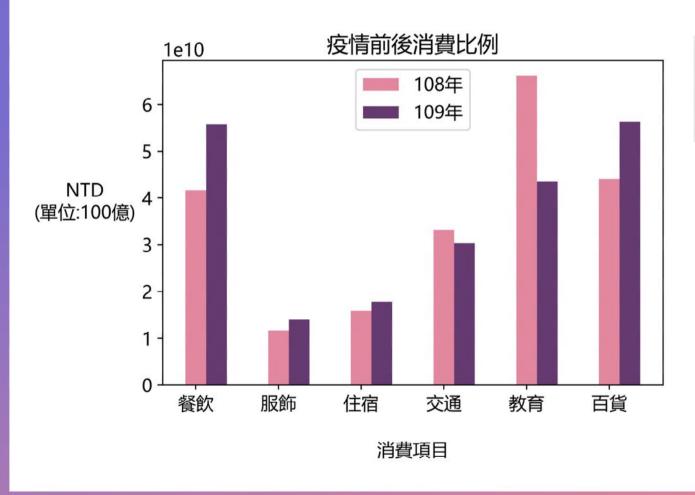
#### 色碼

#643A71
#8B5FBF
#D4C1EC
#FECOCE
#E3879E
#D183C9

- 50萬以下族群的服飾費及交通費的整體比例, 相較200萬以上族群高。
- 200萬以上族群,則會把 錢花費在**住宿**以及**百貨** 購物上。

# п"П

# 02 圖表說明 - 疫情前後 (台北市 50萬以下 108-109年度)



#### 製圖說明

製作有兩個數據的長條圖做比對,找到如何 將109年的數據往右邊移。

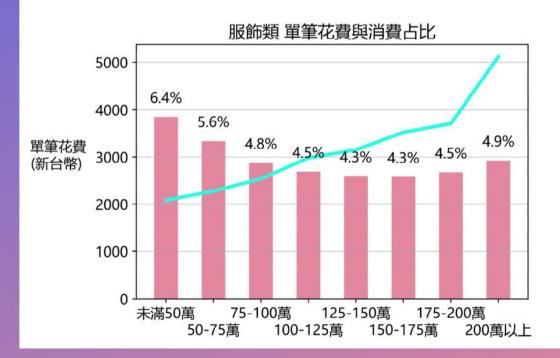
- 看出疫情後,50萬以下族群的**餐飲費**大 幅提高,推估可能**隔離點外送服務**有關。
- 疫情後文教康樂類消費大幅減少,推 估可能由於避免群聚,109年度取消很多 大型的展覽聚會,以及不願意將錢投資 在教育類上面,轉而都在吃東西以及百 貨購物上。



# 02 圖表說明 - 109年單筆花費與消費占比 (服飾類、百貨類)

#### 製圖說明

- 數據1 折線圖 該項目的平均單筆花費
- 數據2長條圖 該項目的花費與其他所有消費的百分比



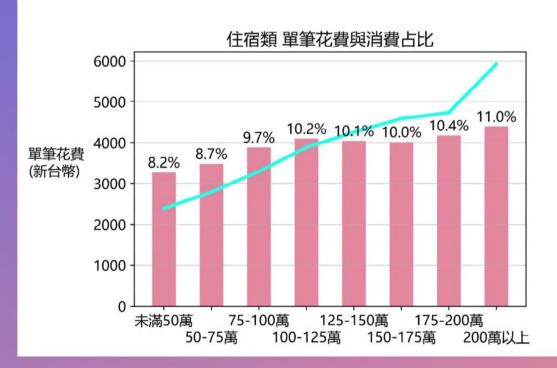
- 50萬以下族群衣服單價在2000元初頭,已是他們消費中的 高百分比了。其他服飾落在3000多就差不多了,200萬年 薪者的衣服都是5000多元。
- 百貨類看出,200萬高年薪者特別喜愛逛百貨購物、單筆消費都在5000元以上。





# 02 圖表說明 - 109年單筆花費與消費占比 (住宿類、交通類)

 越高年薪者,住宿比例越高,顯示有較多時間與金 錢旅遊,而一趟所花費的旅館費用,由50萬以下收 入者住2500元以內的房型開始。



 越高薪,交通費的花費百分比越少,但單筆花費 是高的,推估可能比較少搭大眾交通工具,有機 會搭必定為高鐵或是計程車,或者是由於自駕車 比較高級耗油低,或是有專屬的司機。





Problems and solutions



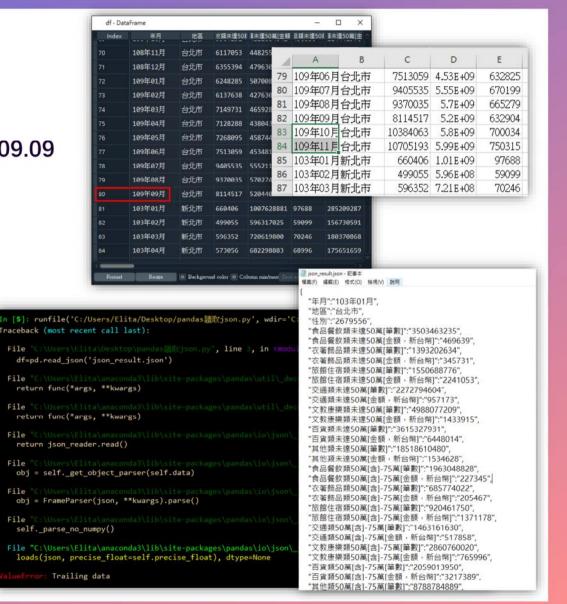


## 03 問題與解決 問題一

#### 問題一 原始檔CSV數據到109.11 JSON檔只到109.09

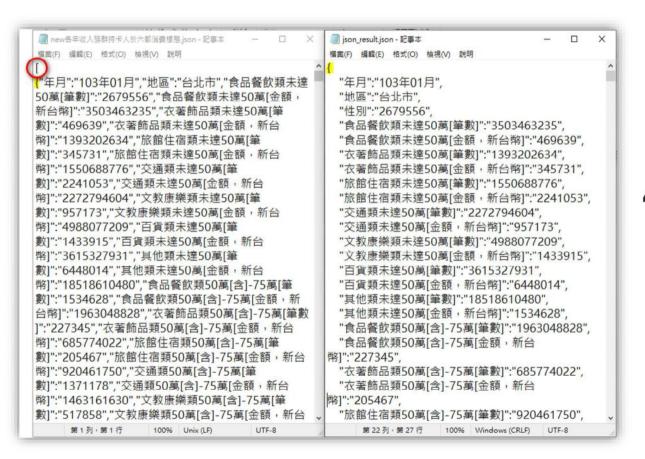
但是, 我只想用JSON檔做資料。

- **1 當試Python轉檔 CSV to JSON** 仿照網路方式寫code轉檔,有轉檔成功但那不是真的JSON檔,spyder無法讀取。
- 2 決定使用線上轉檔工具
  https://www.aconvert.com/tw/document/csv-to-json/
- 3 轉檔成功 得到一份精美的JSON檔!





### 03 問題與解決 問題一



比較與我自己轉檔的差異

左 正確版 / 右 我的失敗版

看到引號的位置不對. 還有我的版本內容有大 量的換行跟空格,之後可再針對這部分做調整。



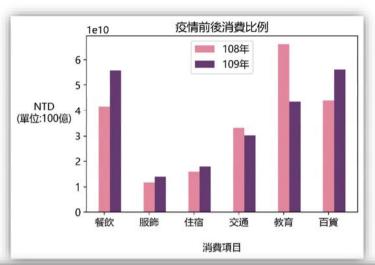


## 03 問題與解決 問題二

#### 問題二 不會兩條數據並排

網路上找到好心人提供解法。

- 2 第二組長條圖等寬右位移 X2的位置就是根據X1的位置往右邊移





## 03 問題與解決 問題三



#### 問題三 如何讓不同數據的折線圖與長條圖並存

單筆花費之數據是千位數、消費占比之數據是個位數。

1 將個位數之數據全部乘600放大

2 最後標籤文字再除600回來

```
#設定資料 服飾類的單筆花費 與 服飾類在所有項目中的占比
single pay=[cloth1,cloth2,cloth3,cloth4,cloth5,cloth6,cloth7,cloth8]
percent=[cloth1_percent*600,cloth2_percent*600,cloth3_percent*600
        ,cloth4 percent*600,cloth5 percent*600,cloth6 percent*600
        ,cloth7 percent*600,cloth8 percent*600]
        #因為百分比數字太小無法跟單筆花費在同一聽表呈現,故將百分比全部乘上600放大數字
plt.figure(dpi=500)
#折線圖 呈現服飾類的單筆花費
plt.plot(x, single_pay, color='#26FFE6', lw=3)
#長條圖 呈現服飾類在所有項目中的占比
plt.bar(range(8),percent,0.5,color='#E3879E')
#Y軸標籤
plt.ylabel('單筆花費 \n(新台幣) ',rotation=0,ha='right')
#圖表標題
plt.title('服飾類 單筆花費與消費占比')
#只要Y軸的網格線
plt.grid(axis='y',linewidth=0.5)
#X軸的刻度標籤 太擠了上下呈現
plt.xticks(range(8),['未滿50萬','\n50-75萬','75-100萬','\n100-125萬','125-150萬
                  ,'\n150-175萬','175-200萬','\n200萬以上'])
#長條圖上方的數值標籤 由於剛剛有乘上600 這邊要除回去
for x,y in enumerate(percent):
   plt.text(x,y+300 '%.1f%%'%(y/600),ha='center')
#顯示圖形
plt.show()
```



Conclusions

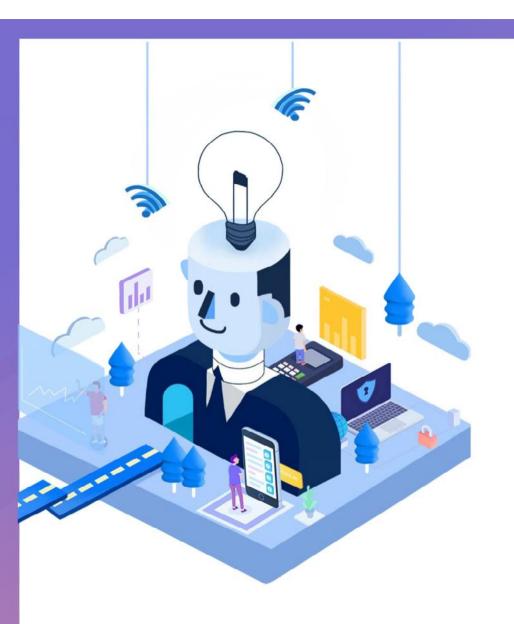


# **)** 1 04 總結與感想

除了各年收入族群持卡人於六都消費樣態

,可以更深入細分**各年收入男性族群**與**各年收入 女性族群**以及**各地區族群**,或是分為**各學歷族 群、各職業族群、各年齡層族群**的消費分析,資 料越豐富分析會越精準,可以協助企業精準找到潛 在消費族群。





# 報告結束

# 感謝各位聆聽

報告者: